



THE AGA KHAN UNIVERSITY

eCommons@AKU

Department of Biological & Biomedical Sciences

Medical College, Pakistan

May 2009

Genetic variations in a well conserved 5-untranslated region of hepatitis C virus genome isolated in Pakistan

Anila Yasmeen
Aga Khan University

Anwar Ali Siddiqui
Aga Khan University, anwar.siddiqui@aku.edu

Saeed Hamidb
Aga Khan University

Taranum Sultana
Aga Khan University

Wasim Jafri
Aga Khan University

See next page for additional authors

Follow this and additional works at: http://ecommons.aku.edu/pakistan_fhs_mc_bbs

 Part of the [Biochemistry, Biophysics, and Structural Biology Commons](#), and the [Medical Sciences Commons](#)

Recommended Citation

Yasmeen, A., Siddiqui, A., Hamidb, S., Sultana, T., Jafri, W., Perssonc, M. A. (2009). Genetic variations in a well conserved 5-untranslated region of hepatitis C virus genome isolated in Pakistan. *Journal of Virological Methods*, 160(1-2), 38-47.
Available at: http://ecommons.aku.edu/pakistan_fhs_mc_bbs/248

Authors

Anila Yasmeeen, Anwar Ali Siddiqui, Saeed Hamidb, Taranum Sultana, Wasim Jafri, and Mats A.A. Perssonc



Genetic variations in a well conserved 5'-untranslated region of hepatitis C virus genome isolated in Pakistan

Anila Yasmeen^a, Anwar Ali Siddiqui^{a,*}, Saeed Hamid^b,
Taranum Sultana^a, Wasim Jafri^b, Mats A.A. Persson^c

^a Department of Biological and Biomedical Sciences, Medical College, Aga Khan University, PO Box 3500, Stadium Road, Karachi 74800, Sindh, Pakistan

^b Department of Medicine, Aga Khan University, PO Box 3500, Stadium Road, Karachi 74800, Pakistan

^c Department of Medicine and Center for Molecular Medicine, L8: 01, Karolinska Institutet, Stockholm, Sweden

A B S T R A C T

Article history:

Received 6 September 2008

Received in revised form 13 April 2009

Accepted 20 April 2009

Available online 3 May 2009

Keywords:

Hepatitis C virus

Genotype 3

Pakistan

Phylogenetic analysis

Neighbour-joining tree

The diversity and extent of sequence variations between hepatitis C virus (HCV) isolates from Pakistan were studied and the probable effects of these variations were assessed on secondary viral structures. Sequencing and phylogenetic analysis was performed on 33 samples, of which 25 were typed as genotype 3 by RFLP (restriction fragment length polymorphism) and 8 remained unresolved. Rooted neighbour-joining (NJ) tree revealed that 28 isolates were HCV type 3a and 5 isolates were typed as 3b. The majority of unresolved samples clustered in a different branch of genotype 3, supported by a bootstrap value of 71%. Another, cluster, cluster I, was found to have a bootstrap value of 81%. Genetic distance values showed significant diversity of isolates in these two clusters compared to the reference sequences. Pair-wise comparison showed the presence of additional restriction sites of *HaeIII* and *RsaI* in unresolved isolates. In conclusion, unique sequence variability was observed in the 5'-UTR of HCV type 3 isolates from Pakistan. One of the reasons for this sequence variability is the presence of mutations, which are additional restriction sites in the 5'-UTR. These mutations were also responsible for failure of conventional RFLP to type some of the HCV isolates.

© 2009 Elsevier B.V. All rights reserved.

1. Introduction

Hepatitis C virus (HCV), a hepacivirus member of the *Flaviviridae* family, has a positive-stranded RNA genome, which consists of a single open reading frame (ORF) and untranslated regions (UTRs) at the 5' and 3' ends. Based on phylogenetic analysis, HCV variants discovered to date are clustered into a two-tiered classification; major genotypes are clustered into six clades, and each genotype has several subtypes (de Lamballerie et al., 1997; Mizokami et al., 1996; Robertson et al., 1998; Simmonds et al., 1996). The importance of determining the HCV genotype has increased since several investigators have described a significant correlation between hepatitis C virus genotype and response to interferon treatment or disease severity (Dammacco et al., 2000; Farci and Purcell, 2000).

Abbreviations: 5'-UTR, 5'-untranslated region; HCV, hepatitis C virus; ORF, open reading frame; E1, envelope 1; RFLP, restriction fragment length polymorphism; ALT, alanine amino transferase; RT-PCR, reverse transcriptase polymerase chain reaction; MP, maximum parsimony; NJ, neighbour-joining; IRES, internal ribosomal entry site.

* Corresponding author. Tel.: +92 21 486 4404; fax: +92 21 493 4924/493 2095.

E-mail addresses: anila.yasmeen@aku.edu (A. Yasmeen), anwar.siddiqui@aku.edu (A.A. Siddiqui), saeed.hamid@aku.edu (S. Hamid), taranum.sultana@aku.edu (T. Sultana), wasim.jafri@aku.edu (W. Jafri), mats.persson@cmm.ki.se (M.A.A. Persson).

Although HCV demonstrates a high degree of sequence variability, the levels of heterogeneity differ considerably among various regions of the genome. The untranslated regions are highly conserved between genotypes (Bukh et al., 1992; Kolykhalov et al., 1996), as is the initial coding region (core region). However, the E1 and E2 genes are highly variable and typically differ at over 50% of sites between genotypes. Interestingly, a short sequence of 5'-UTR region has the potential to be used as a novel target for anti-HCV therapy (Ray and Das, 2004).

A number of methods are available for genotyping HCV (Bukh et al., 1995; Nakao et al., 1991; Machida et al., 1992; Dixit et al., 1995). Sequence analysis of the whole HCV genome remains the gold standard, but it is laborious and expensive, as are the immunoblot assays available commercially. RFLP analysis of sequences in the 5'-UTR of the HCV genome (Davidson et al., 1995) is also used widely in the clinical setting and allows the correct identification of the HCV genotype in more than 90% of cases.

HCV is a major cause of chronic hepatitis in Pakistan where genotype 3 has been reported as the most prevalent type (Bukhtiar et al., 2003; Khokhar and Niazi, 2003; Shah et al., 1997). However, data are lacking on the variability of HCV genome in different regions of Pakistan. The present study aimed at an in-depth analysis of sequence variability in a well conserved region i.e. 5'-UTR of HCV isolates from Pakistan. The 5'-UTR was selected because of

its importance as an essential component of the internal ribosome entry site (IRES) that regulates cap-independent translation of HCV (Wang et al., 2000). For the identification of genotype distribution of different isolates from patients infected with HCV at present in Pakistan, RFLP analysis was used (as it is a more economical assay for screening a large number of samples). DNA sequencing of the 5'-UTR of HCV isolates was performed for the analysis of genetic variation among different isolates. Selected samples were assessed further by phylogenetic analysis to confirm the extent of variations in the 5'-UTR of HCV isolates. Pair-wise alignment of the sequences was done in order to determine the exact location of mutations in the sequences of HCV isolates. Additionally, the effect of mutations was observed on the putative secondary structure model of the viral genome.

2. Materials and methods

2.1. Patients

Two hundred patients (116 male and 84 female; with a mean age of 45 years) with chronic hepatitis C virus infection presenting at the Gastroenterology clinics at Aga Khan University Hospital between 1998 and 2001 were included consecutively in the study. All patients had abnormal ALT levels (122.9 ± 72 IU/ml), positive serum HCV antibody as determined by 3rd generation ELISA (Abbott Laboratories, North Chicago, IL, USA) and negative hepatitis B surface antigen. Serum samples were collected prospectively, aliquoted and stored at -70°C for virological tests. The study was approved by the Aga Khan University Ethical Review Committee (ERC) and informed consent was obtained from all patients prior to participation in the study.

2.2. Virological tests

2.2.1. Amplification of HCV 5'-UTR

Total RNA was extracted from serum by using Tri Reagent as per manufacturer's instructions (Sigma, Chemical Co. St Louis, MO, USA). The RNA pellet was reconstituted in DEPC-treated water containing RNase inhibitor and stored at -70°C until further use.

HCV RNA was amplified by RT-PCR followed by a nested PCR using genotyping primers, as described by Chan et al. (1992). For first strand cDNA synthesis, viral RNA was reverse transcribed by using anti-sense primer (#209) (Chan et al., 1992) to the 5'-UTR. Cycling conditions for PCR were essentially similar to the one described by Chan et al. (1992). In outer PCR, cDNA was used as the template to amplify a region of 600 bp. For amplification, cDNA was supplemented with PCR mixture of 20 μl , containing PCR buffer (1 \times), MgCl_2 (1.5 mM), Triton-X 100 (0.1%), dNTPs (200 μM), 200 nM of sense primer and 1 unit of Pfu Taq polymerase (Promega). Target sequence amplification was done by cycling conditions as described (Chan et al., 1992). The amplified product was used as template for inner PCR or stored at -20°C .

Inner PCR was performed under the similar conditions as used for outer PCR. Outer PCR product (2 μl) was used as template with 200 nM of nested primers (940 and 211). Amplified products (250 bp) were analyzed by electrophoresis on 2% agarose gel.

2.3. RFLP analysis for genotype identification

To determine the genotype of HCV, RFLP analysis was done as described by Davidson et al. (1995). Briefly, amplified products were digested by 2 sets of restriction endonucleases, *RsaI* + *HaeIII* and *HinfI* + *MvaI*. Digested products were studied by electrophoresis on 16% polyacrylamide gel. The genotype was determined by the pattern of the bands present in each sample.

2.4. Sequencing

For the sequencing reaction, an amplified fragment of the 5'-UTR (250 bp) was purified and directly sequenced using standard protocols for the ABI 377 automated sequencer (Applied Biosystems, Foster City, CA, USA). Primers used for sequencing were the same as used in the nested PCR (#941 and #211).

2.5. Nucleotide sequence variability in the HCV 5'-UTR

Nucleotide sequence variability was determined in the 5'-UTR of the HCV genome isolated from different patients. Phylogenetic analysis of 5'-UTR sequences was carried out to identify the clade and subtype distribution of different HCV isolates, especially those which were not resolved by the conventional RFLP method.

2.5.1. BLAST analysis

HCV sample sequences were submitted to BLAST programme to identify the similarity of sample sequences with the reported databases. Sequences were submitted to EMBL. The EMBL accession numbers of sequences reported in this paper are **AM228866–AM228898**.

2.5.2. Reference sequences retrieval

At least 10 sequences of each HCV genotype (1–6) were retrieved from GenBank database to construct rooted neighbour-joining (NJ) and maximum parsimony (MP) tree. Accession numbers of sequences are given below.

2.5.2.1. Accession numbers of reference sequences. M62321-1a, AF009606-1a, AF011751-1a, D29816-1a, D31601-1a, Y10150-1a, D50480-1b, AB049087-1b, AB016785-1b, AF054247-1b, AJ132996-1b, D00831-1b, D14853-1c, AB047645-2a, AB047642-2a, AB047641-2a, AB047640-2a, AB047639-2a, AF177037-2a, D10988-2b, AB030907-2b, D31606-2b, D50409-2c, D17763-3a, AF046866-3a, D29819-3a, AJ006318-3a, AJ006323-3a, D28917-3a, D493743b, D166203d, D166183e, D166143f, X914213g, Y11604-4a, Y13184G-5a, AF064490G-5a, D63822-6a, Y12083-6a.

2.5.3. Manual multiple sequence alignment

Sample (HCV isolates) and reference sequences of HCV 5'-UTR were aligned manually by using SE-AL software (<http://evolve.zoo.ox.ac.uk/software>). Secondary structure information regarding 5'-UTR was also considered during manual alignment.

2.5.4. Phylogenetic analysis by PAUP* (phylogenetic analysis using parsimony* and other methods)

2.5.4.1. Phylogenetic tree construction.

2.5.4.1.1. Maximum parsimony and neighbour-joining tree construction. Rooted maximum parsimony and neighbour-joining trees were constructed by using PAUP* software (Sinauer Associates, Inc., Sunderland, MA, USA). Data were subjected to ModelTest (www.modeltest.com) for DNA substitution modeling required for the construction of a phylogenetic tree. Data was evaluated by ModelTest which showed F84 as the best DNA substitution models for further analysis. The DNA substitution model F84 (Felsenstein and Churchill, 1996) was used to construct neighbour-joining tree. Reliability of the NJ and MP trees was evaluated statistically by bootstrap analysis with 1000 replicates.

2.5.5. Average genetic distances of reference and sample sequences

DNADIST software was used to calculate the distance matrix of sample and reference sequences. Genetic distance matrix was used to calculate distances between all possible pairs of sequences.

The following average genetic distances were calculated further: (a) the average distances within reference sequences; (b) the average distances between reference sequences and samples present in cluster; (c) the average distances between reference and samples resolved by RFLP analysis; and (d) the average distances between reference sequences and unresolved samples

2.6. Pair-wise comparison of HCV isolates

HCV samples which were not resolved by the conventional RFLP method were also sequenced and subjected to pair-wise comparison. Gene Jockey software (<http://www.eskimo.com/~pristine/biology.html#genejockey>) was used for pair-wise comparison of HCV sequences. Each HCV sample sequence of a particular geno-

type (2 or 3) was aligned with the reference sequence of the same genotype (2 or 3).

2.7. Secondary structure of the 5'-UTR

HCV isolates showing unique sequences were studied in further detail. Secondary structure of RNA for these isolates was deduced using RNAviz software (<http://rrna.uia.ac.be/rnaviz/>).

2.8. Statistical analysis

All data are expressed as means \pm standard deviation (S.D.). The degree of significance between various parameters of HCV infected patients was tested by Student's *t*-test.

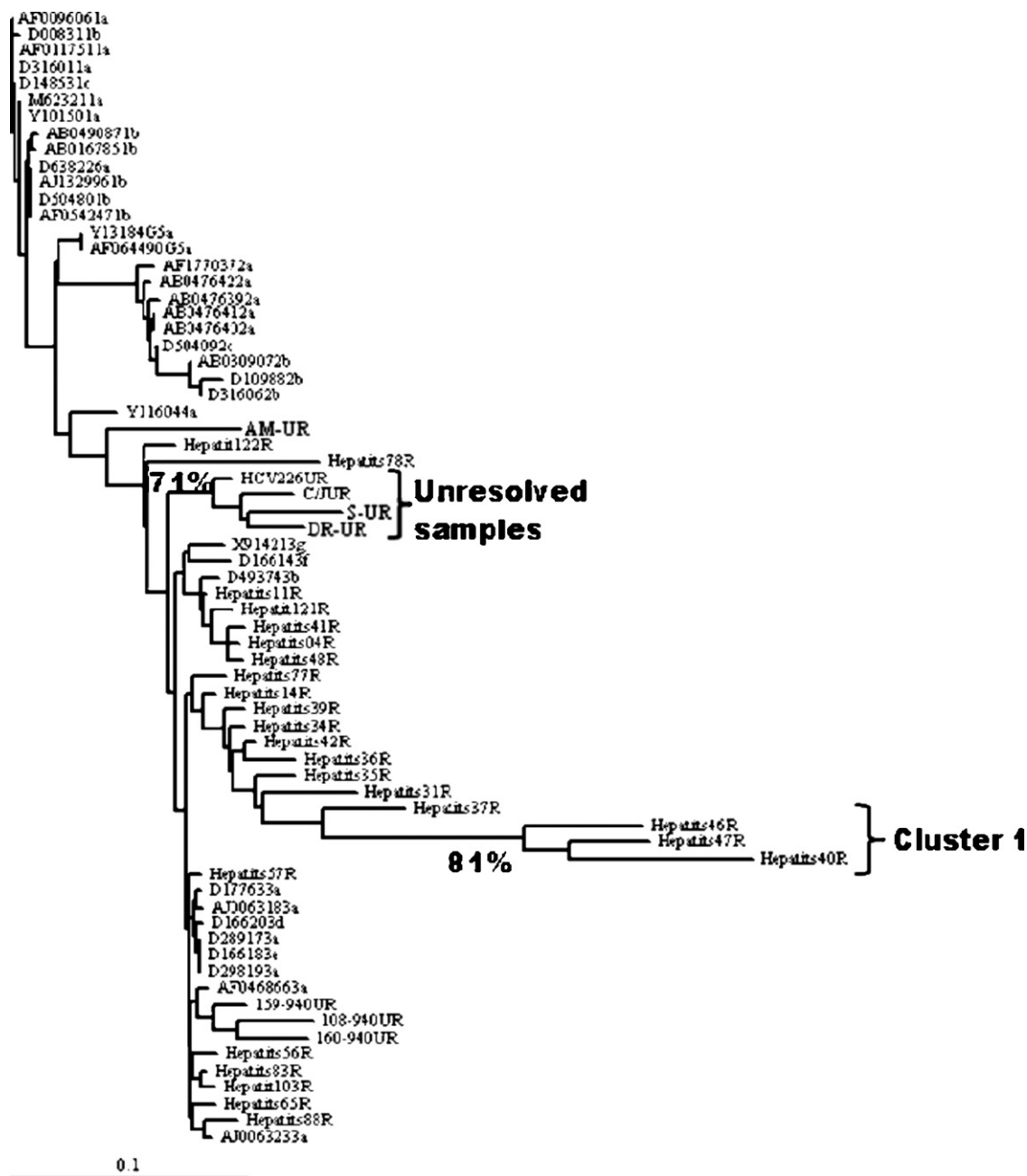


Fig. 1. Rooted NJ tree of reference and HCV sample sequences of 5'-noncoding region. Multiple sequence alignment was performed manually using SE-AL software. Tree indicates the presence of some unusual sequences (cluster 1; bootstrap value 81%) appearing in a separate branch, in addition to unresolved samples (bootstrap value 71%). The reference sequences are indicated by accession numbers and genotype or subtype is indicated in parentheses. Samples of this study are indicated by R (resolved) or UR (unresolved) in parentheses.

3. Results

3.1. RFLP analysis for genotype identification

Out of the 200 samples analyzed, 190 (95%) were found to be positive for HCV RNA. The RFLP results showed that HCV genotype 3 was the most prevalent type (72%), followed by genotype 2 (23%). Only two patients were infected with type 1 (1%). A few samples (4.5%) remained unresolved by RFLP analysis.

3.2. Nucleotide sequence variability in the HCV 5'-UTR

3.2.1. Blast analysis

All unresolved samples (8 HCV isolates) and samples selected at random of genotype 3 (25 HCV isolates) were sequenced for a 250 bp fragment in the 5'-UTR. Blast analysis confirmed that both unresolved and randomly selected samples had high similarity scores with the reported sequences of HCV genotype 3 for which the 5'-UTR region was used.

3.2.2. Phylogenetic tree construction

Sequences of samples included in this study and reference sequences (clades 1–6) were aligned manually while considering the secondary structures present in the 5'-UTR. Aligned sequences were subjected to construction of trees followed by bootstrap analysis. Rooted neighbour-joining tree as depicted in Fig. 1 showed the presence of clades 1–6 as separate genetic lineages. All genotype 3 samples included in this study were present in clade 3. These were, however, separated in different clusters as described below. Robustness of the rooted NJ tree was confirmed by maximum parsimony tree and both trees showed identical results (Figs. 1 and 2).

3.2.2.1. Unique sequences of genotype 3a. It was found that three samples (40R, 46R and 47R) were present in a separate cluster (cluster 1; Fig. 1). This cluster branched out within the major clade of type 3. The reliability of cluster 1 was supported by a bootstrap value of 100% (Fig. 1). Further analysis showed existence of a significant ($p < 0.01$) genetic distance between cluster 1 and reference

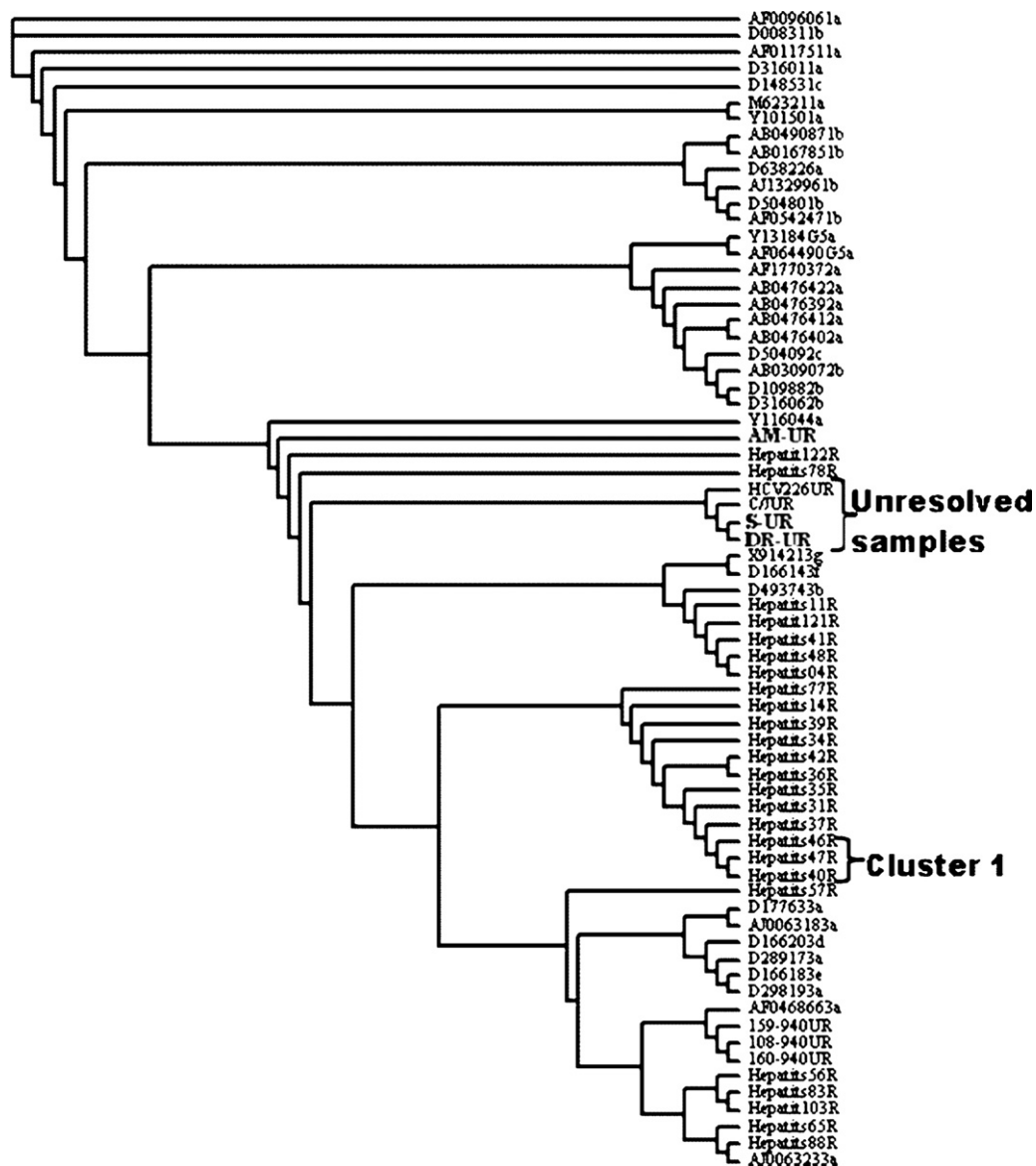


Fig. 2. Maximum parsimony tree of reference and HCV sample sequences of 5'-noncoding region. Multiple sequence alignment was performed manually using SE-AL software. Tree indicates the presence of some unusual sequences (cluster 1) appearing in a separate branch, in addition to unresolved samples in a separate branch. The reference sequences are indicated by accession numbers and genotype or subtype is indicated in parentheses. Samples of this study are indicated by R (resolved) or UR (unresolved) in the parentheses.

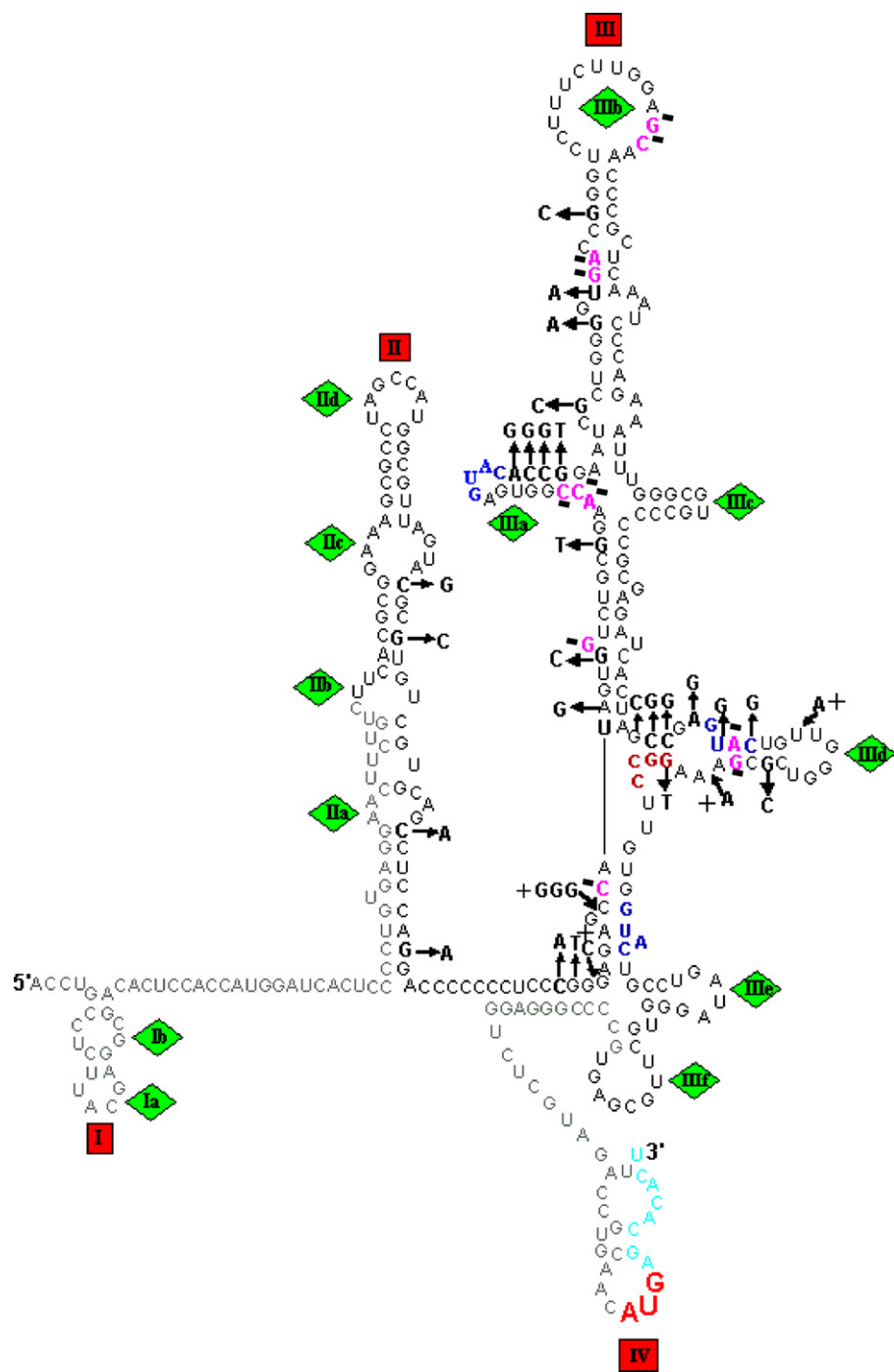


Fig. 3. Nucleotide sequence and putative secondary structure of HCV RNA (nt 1–351, genotype 3a (D17763)). A fragment of 250 bp (61–310 bp) was sequenced (indicated in black letters), while the remaining sequence (grey letters) codes for the remainder of the 5′-UTR. Principal domains are indicated by Roman numerals (I–IV), sub-domains are indicated in diamonds. Substitution sites are indicated as bold and enlarged font, mutations by arrows, insertions by a plus (+) sign and deletions by minus (–) sign.

Table 1
Average genetic distances by using F84 DNA substitution model.

Sequences	Genetic distance (mean ± S.D.)	T-test	p-Value
Reference	0.00816 ± 0.000471		
Cluster 1	0.282677 ± 0.04858	Reference vs. cluster 1	p < 0.01
Unresolved (separate branch)	0.074436 ± 0.024124	Reference vs. unresolved	p < 0.01
Resolved	0.025718 ± 0.011115		
Unresolved	0.057964 ± 0.013576	Resolved vs. unresolved	p < 0.27

Genetic distances calculated by PAUP* using F84 DNA substitution model. Reference sequences include D1776 3a, AF04686 3a, D29819 3a, AJ006318 3a and AJ006323 3a. In cluster 1 46R, 47R and 40R were included. Unresolved (separate branch) sequences include HCV226UR, C/JUR, Salma UR and Dr. Rashid UR. Statistical significance was calculated by Student's *t*-test; *p* < 0.05 was considered significant result.

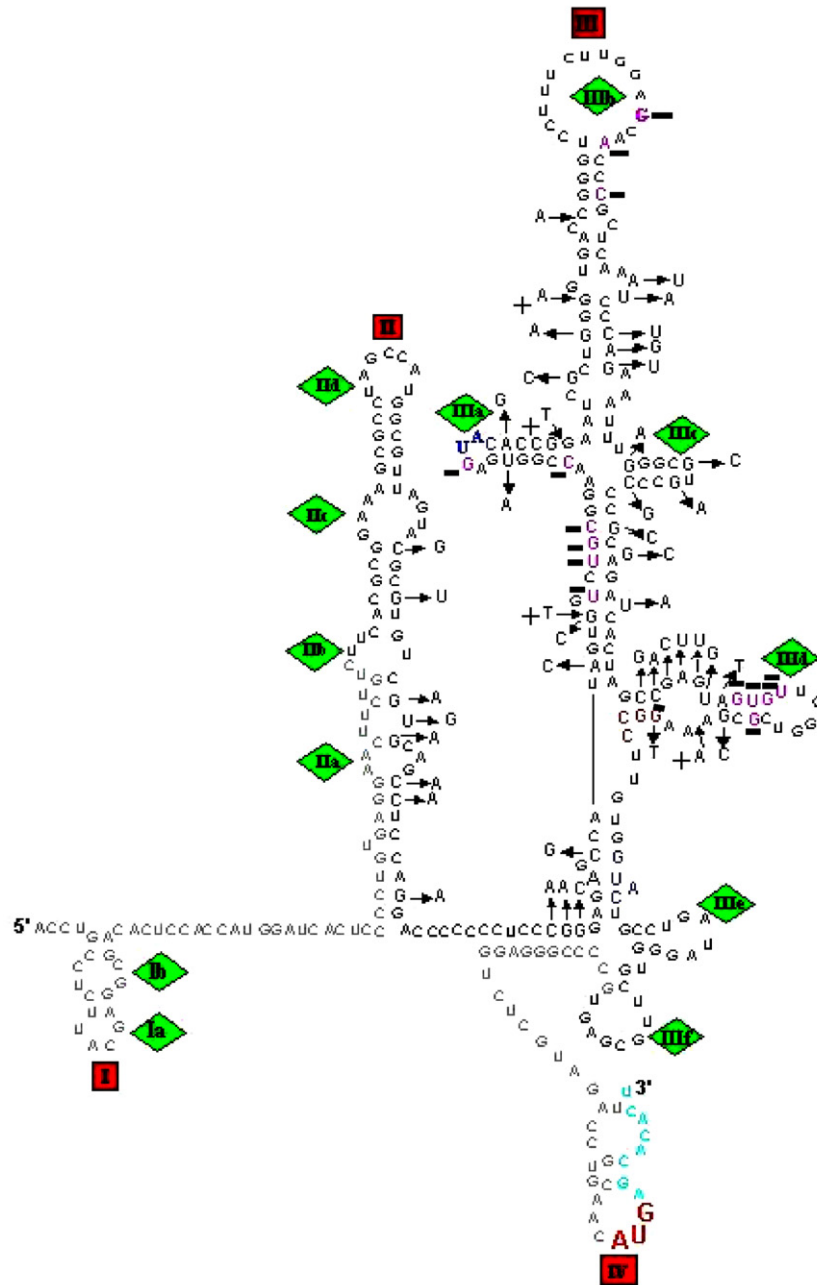


Fig. 4. Nucleotide sequence and putative secondary structure of HCV RNA (nt 1–351, genotype 3a (D17763)). A fragment of 250 bp (61–310 bp) was sequenced (black) while the remaining sequence (grey) codes for the remainder of the 5'-UTR. Principal domains are indicated by Roman numerals (I–IV), sub-domains are indicated in diamonds. Substitution sites are indicated as bold and enlarged font, mutations by arrows, insertions by a plus (+) sign and deletions by minus (–) sign.

sequences (Table 1) despite their identification as genotype 3 by RFLP analysis.

3.2.2.2. Genotype 3b sequences. Five samples (48R, 04R, 41R, 121R and 11R), clustered in a branch as genotype 3b, could not be subtyped further by RFLP analysis (Fig. 1).

3.2.2.3. Unresolved samples by RFLP analysis. Samples that could not be genotyped by conventional RFLP analysis ($n=4$) were also present in the cluster of resolved samples. However, four of these unresolved samples were found in a distinct cluster, supported by a bootstrap value of 71% (Fig. 1). Genetic distances calculated by F84 model showed that these unresolved

samples were present at a significant ($p<0.01$) distance from reference sequences and those that were typed by RFLP method (Table 1).

3.3. Pair-wise comparison of HCV isolates

Pair-wise comparison of HCV sequences showed that additional restriction endonuclease sites (HaeIII and RsaI) were present (Table 2) within the sequences of unresolved samples. Additional nucleotides were also present in unresolved samples, which were clustered in a separate branch. Samples present in cluster 1 did not show any additional site for restriction endonucleases, however, additional stretches of nucleotide were present in these samples.

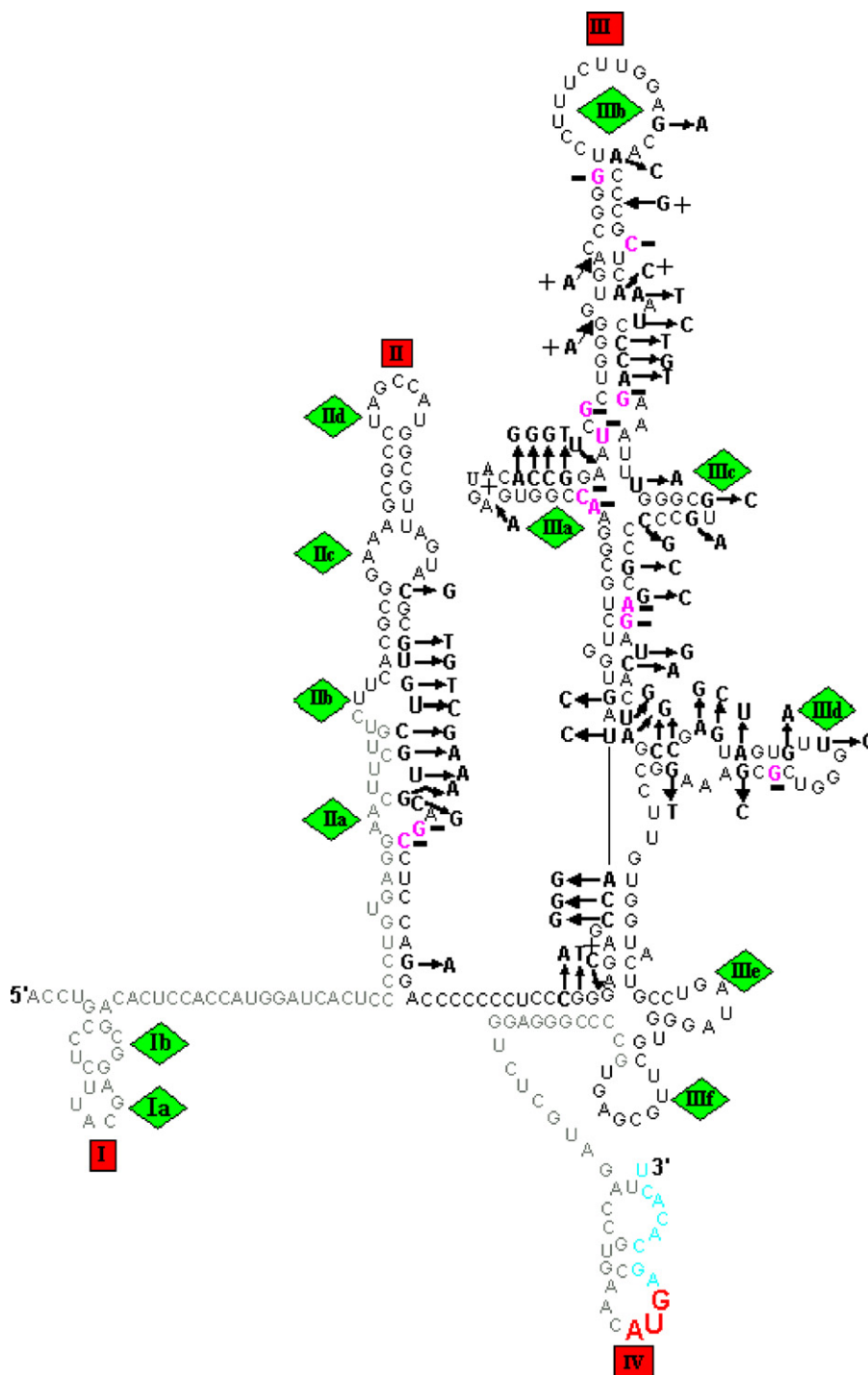


Fig. 5. Nucleotide sequence and putative secondary structure of HCV RNA (nt 1–351, genotype 3a (D17763)). A fragment of 250 bp (61–310 bp) was sequenced (black letter), while the remaining sequence (grey) codes for the remainder of the 5'-UTR. Principal domains are indicated by Roman numerals (I–IV), sub-domains are indicated in green diamonds. Substitution sites are indicated as bold and enlarged font, mutations by arrows, insertions by a plus (+) sign and deletions by minus (–) sign. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of the article)

3.4. Secondary structure of 5'-UTR

Secondary structure of HCV-RNA 5'-UTR (250 bp fragment) was deduced by RNAviz software. Nucleotide substitutions or deletions found in samples present in cluster 1 were analyzed for their putative effect on the secondary structure of this region. Substantial disruption was noted in stem loops of this region. It was found that mutations were mainly concentrated in stem loop IIIId of the 5'-UTR region (Figs. 3–5). In sample

#47R (Fig. 5), the genetic distance was greater than other samples as indicated by mutations in loops IIIId, IIIb and also in domain II.

Samples that were not resolved by RFLP (showing additional restriction sites) were subjected to analysis of their secondary structure. Except for one sample (#AM-UR), minimal effect on secondary structures were observed (Fig. 6). In sample #AM-UR, an additional HaeIII restriction site resulted in addition of another loop next to loop IIIb.

initiation function of the virus as reported previously (Yasmeen et al., 2006).

4.1. Unique mutations in the 5'-UTR sequences

A notable variation was observed between various sequences of type 3a in HCV isolates. Sequences falling into the 3rd clade separated from each other in distinct clusters (cluster 1; Fig. 1) (Chan et al., 1992). Branch lengths for HCV isolates in cluster 1 are considered to indicate a more rapid evolution process than the average of the other major genotypes. The average genetic distance of these sequences was also significantly different from the reference sequences.

Although the possibility of the presence of a new subtype can be inferred from these findings, it is considered that such observations need to be confirmed and substantiated by further investigations using more than one approach. Therefore, this particular finding is not assigned as a new genotype. It is recognized that assigning these unique sequences a new subtype will require an in-depth exploration on sequence variability in at least two coding regions (Simmonds et al., 1994). However, the present information on phylogenetic relatedness provides a more appropriate description of the evolutionary and epidemiological history of a virus. Recombination with some other viruses is another possibility that may provide an answer to the alterations found in these unique sequences.

4.2. Unresolved samples

Blast analysis and phylogenetic analysis also revealed the genotype of samples that were not typeable by using conventional RFLP analysis. The existence of genetic variants containing additional restriction sites may affect the RFLP results, which is otherwise a simple and cost effective method for clade identification of HCV in screening of large number of samples. In clinical practice, correct virus genotyping has been shown to be critical for determining optimal treatment duration (Poynard et al., 1998).

4.3. Structural implications of 5'-UTR sequence variation

The findings of covariant sequence change within the 5'-UTR coupled with the location and measure of other reported sites of polymorphism provide good evidence in support of the secondary structure model for HCV 5'-UTR proposed earlier (Brown et al., 1992) and modified here.

Nucleotide substitution within the HCV 5'-UTR may influence the viral translation and its sensitivity to the antiviral action of interferon. It has been reported that domain II induces a conformational change in the 40S ribosomal subunit, which has been implicated in the RNA decoding process, whereas domain III has been shown to interact with 40S ribosomal subunit and some trans-acting factors critical for IRES-mediated translation (Ali and Siddiqui, 1997; Kieft et al., 2001; Kolupaeva et al., 2000). Mutations in this domain disrupt IRES-mediated translation initiation (Psaridi et al., 1999). In the present study, it was found that samples present in cluster 1 showed mutations mainly in stem loop IIIId (Figs. 3–5), which may cause structural reorganization of the HCV IRES and ultimately reduce its activity (Jubin et al., 2000).

With the availability of this new information and the findings related to correlation between sequence changes in the IRES region and translation efficiency of the virus isolated from different patients (Yasmeen et al., 2006), a more rational approach would be to observe the variability in the whole IRES region, including the pseudoknot structure. Functional aspects of these secondary structure models could be confirmed by experiments studying the IRES function of artificial mutants (Wang et al., 1994) or interference anti-sense oligonucleotides (Wakita and Wands, 1994).

5. Conclusions

In conclusion, this report describes the presence of unique sequence variability in the 5'-UTR of hepatitis C virus genotype 3 isolates from Pakistan. Mutations found in these unique sequences could have led to conformational changes in stem loop III that might alter the translation initiation of the virus. In future, major efforts will be focused to analyze the functional impact of these sequences with the inclusion of the whole IRES region and to determine their clinical impact. In the present study, additional restriction sites were identified that result in the failure of conventional RFLP analysis to identify the genotype in some cases. This necessitates sequencing of these isolates to explore the extent of variations at the viral genomic level.

Acknowledgements

We acknowledge the valuable support of Dr. Thomas Leitner in terms of suggestions, guidance to A.Y. and critical reading of the manuscript. We would like to thank Dr. Kulsoom Ghias for critical reading of the manuscript. This work was supported by the University Research Council, AKU and Swedish Institute fellowship awarded to A.Y.

References

- Ali, N., Siddiqui, A., 1997. The La antigen binds 5'-noncoding region of the hepatitis C virus RNA in the context of the initiator AUG codon and stimulates internal ribosome entry site mediated translation. *Proc. Natl. Acad. Sci. U.S.A.* 18, 2249–2254.
- Brown, E.A., Zhang, H., Ping, H.-L., Lemon, S.M., 1992. Secondary structure of the 5'-nontranslated region of hepatitis C virus and pestivirus genomic RNAs. *Nucleic Acid Res.* 20, 5041–5045.
- Bukh, J., Purcell, R.H., Miller, M., 1992. Sequence analysis of the 5' noncoding region of hepatitis C virus. *Proc. Natl. Acad. Sci. U.S.A.* 89, 4042–4046.
- Bukh, J., Miller, R.H., Purcell, R.H., 1995. Genetic heterogeneity of hepatitis C virus: quasispecies and genotypes. *Semin. Liver Dis.* 15 (1), 41–63.
- Bukhtiar, N., Hussain, T., Iqbal, M., Malik, A.M., Quresh, I.A.H., Hussain, A., 2003. Hepatitis B and C single and co-infection in chronic liver disease and their effect on the disease pattern. *J. Pak. Med. Assoc.* 53, 136–140.
- Chan, S.W., McOmish, F., Holmes, E.C., Dow, B., Peutherer, J.F., Follet, E., Yap, P.L., Simmonds, P., 1992. Analysis of a new hepatitis C virus type and its phylogenetic relationship to existing variants. *J. Gen. Virol.* 73, 1131–1141.
- Dammacco, F., Sansonno, D., Piccoli, C., Racanelli, V., D'Amore, F.P., Lauletta, G., 2000. The lymphoid system in hepatitis C virus infection: autoimmunity, mixed cryoglobulinemia, and overt B-cell malignancy. *Semin. Liver Dis.* 20 (2), 143–157.
- Davidson, F., Simmonds, P., Ferguson, J.C., 1995. Survey of major genotypes and subtypes of hepatitis C virus using RFLP of sequences amplified from 5-NCR. *J. Gen. Virol.* 76, 1197–1204.
- de Lamballerie, X., Charrel, R.N., Attoui, H., De Micco, P., 1997. Classification of hepatitis C virus variants in six major types based on analysis of the envelope 1 and nonstructural 5B genome regions and complete polyprotein sequences. *J. Gen. Virol.* 78 (Pt 1), 45–51.
- Dixit, V., Quan, S., Martin, P., Larson, D., Brezina, M., DiNello, R., Sra, K., Lau, J.Y.N., Chein, D., Kolberg, J., Tagger, A., Davis, G., Polito, A., Gitnicki, G., 1995. Evaluation of a novel serotyping system for hepatitis C virus: strong correlation with standard genotyping methodologies. *J. Clin. Microbiol.* 33 (11), 2978–2983.
- Farci, P., Purcell, R.H., 2000. Clinical significance of hepatitis C virus genotypes and quasispecies. *Semin. Liver Dis.* 20, 103–126.
- Felsenstein, J., Churchill, G., 1996. A hidden markov model approach to variation among sites in rate of evolution. *Mol. Biol. Evol.* 13, 93–104.
- Jubin, R., Vantuno, N.E., Kieft, J.S., Murray, M.G., Doudna, J.A., Lau, J.Y., Baroudy, B.M., 2000. Hepatitis C virus internal ribosome entry site (IRES) stem loop IIIId contains a phylogenetically conserved GGG triplet essential for translation and IRES folding. *J. Virol.* 74, 10430–10437.
- Khokhar, N., Niazi, S.A., 2003. Chronic liver disease related mortality pattern in Northern Pakistan. *J. Coll. Physicians Surg. Pak.* 13, 495–497.
- Kieft, J.S., Zhou, K., Jubin, R., Doudna, J.A., 2001. Mechanism of ribosome recruitment by hepatitis C IRES RNA. *RNA* 7, 194–206.
- Kolupaeva, V.G., Pestova, T.V., Hellen, C.U.T., 2000. An enzymatic footprinting analysis of the interaction of 40S ribosomal subunits with the internal ribosomal entry site of hepatitis C virus. *J. Virol.* 74, 6250–6262.
- Kolykhalov, A.A., Feinstone, S.M., Rice, C.M., 1996. Identification of highly conserved sequence element at the 3' terminus of hepatitis C virus genome. *J. Virol.* 70, 3363–3371.
- Machida, A., Ohnuma, H., Tsuda, F., Muneakata, E., Tanaka, T., Akahane, Y., Okamoto, H., Mishiro, S., 1992. Two distinct subtypes of hepatitis C virus defined by antibodies directed to the putative core protein. *Hepatology* 16 (4), 886–891.

- Mizokami, M., Gojobori, T., Ohba, K., Ikeo, K., Ge, X.M., Ohno, T., Orito, E., Lau, J.Y., 1996. Hepatitis C virus types 7, 8 and 9 should be classified as type 6 subtypes. *J. Hepatol.* 24, 622–624.
- Nakao, T., Enomoto, N., Takada, N., Takada, A., Date, T., 1991. Typing of hepatitis C virus genomes by restriction fragment length polymorphism. *J. Gen. Virol.* 72 (Pt 9), 2105–2112.
- Poynard, T., Marcellin, P., Lee, S.S., Niederau, C., Minuk, G.S., Ideo, G., Bain, V., Heathcote, J., Zeuzem, S., Trepo, C., Albrecht, J., 1998. Randomised trial of interferon alpha2b plus ribavirin for 48 weeks or for 24 weeks versus interferon alpha2b plus placebo for 48 weeks for treatment of chronic infection with hepatitis C virus. International Hepatitis Interventional Therapy Group (IHIT). *Lancet* 352 (9138), 1426–1432.
- Psaridi, L., Georgopoulou, U., Varaklioti, A., Mavromara, P., 1999. Mutational analysis of a conserved tetraloop in the 5' untranslated region of hepatitis C virus identifies a novel RNA element essential for the internal ribosome entry site function. *FEBS Lett.* 453, 49–53.
- Ray, P.S., Das, S., 2004. Inhibition of hepatitis C virus IRES-mediated translation by small RNAs analogous to stem-loop structures of the 5'-untranslated region. *Nucleic Acids Res.* 32, 1678–1687.
- Robertson, B., Myers, G., Howard, C., et al., 1998. Classification, nomenclature, and database development for hepatitis C virus (HCV) and related viruses: proposals for standardization. *Arch. Virol.* 143, 2493–2503.
- Shah, H.A., Jafri, W., Malik, I., Prescott, L., Simmonds, P., 1997. Hepatitis C virus (HCV) genotypes and chronic liver disease in Pakistan. *J. Gastroenterol. Hepatol.* 12, 758–761.
- Simmonds, P., Mellor, J., Sakuldamrongpanich, T., Nuchaprayoon, C., Tanprasert, S., Holmes, E.C., Smith, D.B., 1996. Evolutionary analysis of variants of hepatitis C virus found in South-East Asia: comparison with classifications based upon sequence similarity. *J. Gen. Virol.* 77 (Pt 12), 3013–3024.
- Simmonds, P., Smith, D.B., McOmish, F., Yap, P.L., Kolberg, J., Urdea, M.S., Holmes, E.C., 1994. Identification of genotypes of hepatitis C virus by sequence comparisons in the core, E1 and NS-5 regions. *J. Gen. Virol.* 75, 1053–1061.
- Wakita, T., Wands, J.R., 1994. Specific inhibition of hepatitis C virus expression by antisense oligonucleotides-in vitro model for selection of target sequence. *J. Biol. Chem.* 269, 14205–14210.
- Wang, T.H., Rijnbrand, R.C., Lemon, S.M., 2000. Core protein-coding sequence, but not core protein, modulates the efficiency of cap-independent translation directed by the internal ribosome entry site of hepatitis C virus. *J. Virol.* 74 (23), 11347–11358.
- Wang, C.Y., Sarnow, P., Siddiqui, A., 1994. A conserved helical element is essential for internal initiation of translation of hepatitis C virus RNA. *J. Virol.* 68, 7301–7307.
- Yasmeeen, A., Hamid, S., Granath, F.N., Lindström, H., Elliott, R.M., Siddiqui, A.A., Persson, M.A.A., 2006. Correlation between translation efficiency and outcome of combination therapy in chronic hepatitis C of genotype 3. *J. Viral. Hep.* 13, 87–95.